DIGITAL AUDIO WATERMARKING USING SPREAD SPECTRUM AND A PSYCHOACOUSTIC MODEL

BY

OZLEM KALINLI

Submitted in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering
in the Graduate College of the
Illinois Institute of Technology

Approved ______________________________
Advisor

Chicago, Illinois
May 2003

ACKNOWLEDGMENT

I would like to express my sincere thanks to my advisor, Prof. Joseph L. LoCicero, for his guidance, nurturing, and support during my graduate studies. Without his effort this thesis would not be fruitful. His knowledge, kindness, openmindedness, and vision have provided me with lifetime benefits. I enjoyed working with him very much, and I learned a lot from him.

My appreciation goes to Prof. Jafar Saniie, who gave me the opportunity to study at I.I.T.. Working as a T.A. with him was a great experience.

I would like to acknowledge, Deniz Pazarci, my undergraduate inspiration at Istanbul Technical University (I.T.U.). Her faith, encouragement, and love was familial. Also, I am thankful to Prof. Melih Pazarci at I.T.U. for his guidance and his valuable advice.

There are number of friends that I would like to thank. Tarun Tandon, my project partner at the beginning of this research, deserves credit for his insightful discussions on audio watermarking. I appreciate his valuable contributions, and also his friendship. I would like to acknowledge the support and friendship of Fernando Martinez Vallina, who provided a helping hand for the past two years, even when I was sick. Also, I want to thank Erdal Oruklu, and Rifat Hacioglu for their help, and friendship. My sincere thanks goes to my best friend, Elif Güzel, for her understanding, support, and sisterhood for the past ten years.

The love, support, and encouragement of my family cannot be acknowledged enough. My eternal gratitude goes to my parents, Neclâ and Mustafa Kalinli, and my brother Özgur Kalinli. I dedicate this thesis to them.

TABLE OF CONTENTS

ABSTRACT

Digital audio watermarking is a technique developed to help in copyright protection of audio files. The research in this area has been spurred by the growth of the Internet, and the ease of copying audio material. This has led to an unprecedented numbers in copyright infringements. One area of research in digital audio watermarking is concerned with embedding data in an audio signal such that it is inaudible and the watermark data can be accurately detected.

In this thesis, the implementation of digital audio watermarking is studied in the context of baseband (BB) direct sequence spread spectrum (DS-SS), and amplitude modulated (AM) DS-SS with a binary watermark data stream. Performance of these techniques are accessed with audibility tests and the accuracy of watermark data detection. Hard and soft decision detection along with temporal diversity are used to recover the watermark signal. The AM DS-SS method clearly outperforms the BB DS-SS technique.

A Gaussian monopulse (GMP) signal is introduced as an alternative to the normal rectangular spread spectrum waveform. The GMP method produces a bandpass watermark signal that is spread over the frequency band of the wide band audio signal without the need for carrier modulation. Extending the GMP method with the Hilbert transform of the GMP, a novel method called quadrature GMP (QGMP), is proposed. The QGMP method doubles the length of the watermark sequence for a fixed length audio signal, or allows for a doubling of diversity. Performance of the GMP method and the QGMP method with listening tests and watermark detection shows an improvement over the AM DS-SS technique.

Psychoacoustic modelling (PAM) can be used to embed the watermark signal into the audio signal in the subband frequency domain. This allows excellent wa-

CHAPTER VII
CONCLUSION

7.1 Summary

Digital audio watermarking has been investigated in this thesis. Spread spectrum watermarking in baseband (BB DS-SS) and amplitude modulated spread spectrum (AM DS-SS) watermarking are two techniques that were investigated. In these systems, the watermark signal is simply attenuated with a constant before it is added to the audio signal to make it inaudible. The BB DS-SS method and the AM DS-SS method do not require access to the original audio signal to recover the watermark at the receiver. Audibility and watermark detection performance of AM DS-SS is superior to that of BB DS-SS. A post processing technique is proposed for the AM DS-SS method. Watermark detection can be improved by applying a LPF to the demodulated signal. The LPF eliminates audio components that fall into the high frequency band after demodulation of the AM DS-SS watermark signal.

A novel method using Gaussian monopulse waveforms (GMP) is proposed as an alternative to the AM DS-SS method. The GMP method uses the GMP waveform to directly produce a bandpass watermark signal. The GMP method is less complex than the AM DS-SS method. No modulation, demodulation, or filtering is required with the GMP watermarking system. Audibility and watermark detection performance of the GMP method is comparable to that realized with the AM DS-SS method. Watermark recovery in the GMP method is also accomplished without knowledge of the original audio signal at the receiver, as with the AM DS-SS and BB DS-SS methods.

The GMP method is extended to the QGMP method, using the GMP and its Hilbert transform as orthogonal basis functions. This method doubles the length of the watermark sequence for a fixed length audio signal. Alternatively, the temporal diversity can be doubled with the QGMP method while keeping the watermark length the same. Doubled temporal diversity improves watermark detection at the receiver. An eight-phase GMP method is introduced following the QGMP concept. The 8-PGMP triples the watermark length for a fixed length audio signal, or triples the temporal diversity when the watermark length is unchanged. However, watermark detection suffers due to a reduced Euclidean distance.

To make the watermark signal robust to lossy compression techniques and make it more inaudible, the psychoacoustic model (PAM) of MPEG-1 Layer I is used to shape the watermark signal in the frequency domain. The novelty introduced in this thesis is shaping and combining of the audio and watermark in the subband frequency domain. The proposed method shapes the watermark signal in such a way that most of the watermark energy is embedded in the inaudible regions of the audio signal. Also, the audio signal is eliminated in the inaudible regions before the watermark signal is added. Watermarking using a PAM is implemented for the AM DS-SS method, the GMP method, and the QGMP method. The watermarked audio signals are transparent, and classified as being inaudible. Also, the watermark signal can be recovered error free without knowledge of the original audio signal.

A post PAM filtering technique is proposed that improves the watermarking methods. The received signal is applied to the PAM, and the audible regions of the watermarked audio signal are removed in the subband frequency domain using a receiver filter. The filtered signal includes less audio signal noise, and yileds even better watermark detection performance. Again watermark distortion is inaudible.

7.2 Future Research Work

During our digital audio watermarking studies, several items were identified that require further investigation. These are detailed below.

The psychoacoustic model of MPEG-1 Layer I does not generate accurate masking threshold levels, especially for intervals if silence in the audio signal. A more accurate psychoacoustic model can be used to improve both audibility and watermark detection performance. The most promising PAM is the MPEG-1 Layer III algorithm [19].

Eight-phase GMP is introduced in this thesis, but it is not implemented. It is a method that will triple the length of the watermark sequence. Alternatively, the temporal diversity can be tripled. It has the potential to improve watermark recovery. However, coding will likely be needed to decrease the probability of the error due to the eight phase constellation. A combined modulation and coding technique, similar to trellis coded modulation needs to be studied.

There is an entire area in digital audio watermarking that can be investigated using the embedding techniques studied in this thesis. That area is robustness to intentional or unintentional attacks, such as lossy encoding/decoding, and filtering.

For some specific applications, it is necessary to decrease the probability of the false alarm. For these applications, a Bayesian-type of detector can be used. Generally, it is assumed that the unwatermarked contents are uncorrelated. However, this assumption is actually not correct. A *whitening filter* can be used before the detector to decorrelate the samples of the content, and this has the potential to decrease the error rates.